Trimerization of the Heat Shock Transcription Factor by a Triple-Stranded α-Helical Coiled-Coil†

Ralph Peteranderl and Hillary C. M. Nelson*

Department of Molecular and Cell Biology, Division of Biochemistry and Molecular Biology, University of California, Berkeley, California 94720

Received August 4, 1992; Revised Manuscript Received October 2, 1992

ABSTRACT: We have isolated and characterized a 91 amino acid fragment of the heat shock transcription factor from both *Saccharomyces cerevisiae* and *Kluyveromyces lactis*. The two protein fragments behave similarly: they form homotrimers, as indicated by sedimentation equilibrium and cross-linking, and contain approximately 80% α-helix, as indicated by circular dichroism. Sedimentation velocity and diffusion coefficients indicate that they have an elongated, nonspherical shape. We conclude the following: these fragments contain a domain which forms a trimer via a triple-stranded α-helical coiled-coil, similar to that found in influenza hemagglutinin.

The heat shock transcription factor (HSF) is a constitutively expressed transcriptional activator involved in the eukaryotic heat shock response. HSF binds to the heat shock element (HSE), a conserved sequence found upstream of all genes which encode heat shock proteins. The HSE can be described as a series of nGAAn boxes, spaced in alternating orientation every five base pairs. Typically, at least three nGAAn boxes, e.g., nGAAnnTTCnnGAAn, are needed for transcriptional activation in vivo (Xiao & Lis, 1988; Amin et al., 1988). The genes encoding HSF have been cloned from *Saccharomyces cerevisiae*, *Kluyveromyces lactis*, *Drosophila melanogaster*, mouse, human, and tomato (Sorger & Pelham, 1988; Wiederrecht et al., 1988; Jakobsen & Pelham, 1991; Clos et al., 1990; Sarge et al., 1991; Rabindran et al., 1991; Schuetz et al., 1991; Scharf et al., 1990). Initial sequence comparison and deletion analyses indicated the existence of two highly conserved functional domains of the protein: a DNA binding domain and a trimerization domain. The putative DNA binding domain shows no homology to known DNA binding proteins and may therefore be a novel DNA binding motif. The putative trimerization domain is characterized primarily by at least six repeats of a heptad array containing hydrophobic amino acids occupying positions "a" and "d" of the heptad sequence abcdefg. The fact that the protein forms a homotrimer is unique for a DNA binding protein and may be important for recognition of the HSEs (Sorger & Nelson, 1989; Perisic et al., 1989).

Previously, it was shown that fragments of the 833 amino acid *S. cerevisiae* HSF are trimeric in solution and bind to DNA as a trimer (Sorger & Nelson, 1989). Cross-linking results with full-length HSF purified from either *S. cerevisiae* (Sorger, unpublished results) or *D. melanogaster* (Perisic et al., 1989) also confirmed the trimeric nature of the protein. The region necessary for trimerization was localized using C-terminal truncations and fusion proteins (Sorger & Nelson, 1989). C-terminal truncated fragments of the *S. cerevisiae* HSF were cross-linked with ethylene glycol bis(succinimidyl succinate) (EGS). Proteins that contained residues 1–424 were able to form trimeric complexes in these experiments, while proteins that contained residues 1–327 or 1–393 were monomeric. The approximate N-terminal boundary of this domain was defined by the expression of fusion proteins between a bacterial protein (LexA) and truncated versions of HSF. Fusions of LexA with HSF residues 211–424 or 327–424 were able to form heterologous complexes with a truncated version of HSF containing residues 1–424. Taken together, these results predict that residues 327–424 contain the region responsible for oligomerization. The six heptad arrays of hydrophobic amino acids are found within this region at residues 344–385, and are typical of those found in α-helical coiled-coil proteins (Crick, 1953; Hodges et al., 1972; Sodek et al., 1972). Subsequent sequence analysis shows that a very similar pattern of contiguous repeats occurs in all other HSFs. Sorger and Nelson (1989) proposed that the putative trimerization domain of HSF forms a triple-stranded α-helical coiled-coil.

To test these proposals, we have overexpressed residues 333–424 from the *S. cerevisiae* HSF protein and the analogous fragment from the HSF protein of *K. lactis* (residues 304–394) in *Escherichia coli*. The physicochemical properties of these two protein fragments were studied by circular dichroism, cross-linking, analytical ultracentrifugation, and dynamic light scattering. We show that these fragments are not only necessary but also sufficient for trimerization. They are predominantly α-helical and have solution properties consistent with a triple-stranded coiled-coil, similar to that found in influenza hemagglutinin.

MATERIALS AND METHODS

Overexpression and Purification. A fragment of the *S. cerevisiae* HSF gene encoding amino acids 333–424 and an analogous fragment from the *K. lactis* HSF gene encoding amino acids 304–394 were cloned into a derivative of the overexpression vector pET3-b (Studier et al., 1990). A natural methionine at amino acid 333 (*S. cerevisiae*) or a methionine introduced by site-directed mutagenesis at amino acid 303 (*K. lactis*) was used as the start site of translation. The cloning procedure added four amino acids, GMLN, to the C-terminal end of both peptides. The resultant proteins are called Sc-t96 (*S. cerevisiae*) and Kl-t96 (*K. lactis*). The overexpression vectors were transformed into BL21(DE3)/pLysS and induced as described (Studier et al., 1990). Cells were stored as a

† This work was supported in part by U.S. Public Health Service Grant GM-44086 and by a grant from the Lucille P. Markey Charitable Trusts. R.P. was supported by a Florence M. Lankdershim scholarship.

* Address correspondence to this author at Stanley/Donner ASU, University of California, 229 Stanley Hall, Berkeley, CA 94720.

0006-2960/92/0431-12272$03.00/0

frozen suspension in 50 mM Hepes, pH 8.0, 600 mM NaCl, 2 mM EDTA, 5 mM $CaCl_2$, and 10% glycerol.

The purification protocol was identical for both Sc-t96 and Kl-t96. Frozen cell suspensions were thawed, sonicated, and centrifuged at 31000*g* for 30 min at 10 °C. Proteins from the high-speed supernatant were precipitated at 4 °C by the addition of saturated $(NH_4)_2SO_4$ to 30% (v/v), and the pellet was resuspended in 50 mM Hepes, pH 8.0, 600 mM NaCl, 2 mM EDTA, 5 mM $CaCl_2$, and 10% glycerol. After a 1 h dialysis against the same buffer, the dialysate was adjusted to 50 mM Hepes, pH 8.0, 400 mM NaCl, and 800 mM $(NH_4)_2SO_4$. All subsequent chromatographic steps for the purification were performed at room temperature on a Waters 650 Advanced Protein Purification System, using slurry-packed Waters Advanced Purification glass columns. The dialysate was loaded onto a hydrophobic interaction chromatography column (100 × 25 mm i.d., SynChropak Propyl, 300-Å pore size). The column was developed with a decreasing linear $(NH_4)_2SO_4$ gradient at a change of 20 mM $(NH_4)_2SO_4$/min and a flow rate of 4 mL/min. The protein eluted between 400 and 180 mM $(NH_4)_2SO_4$. The column fractions were dialyzed against 50 mM Tris, pH 6.7, 50 mM NaCl, 2 mM EDTA, 5 mM $CaCl_2$, and 10% glycerol, and then loaded at 2 mL/min onto a cation-exchange column (100 × 10 mm i.d., Waters Accell Plus CM, 500-Å pore size, 50-μm average particle size). The protein did not bind to this column, but contaminants did and were separated. The fractions that flowed through the column were dialyzed against 50 mM Tris, pH 8.8, 50 mM NaCl, 2 mM EDTA, 5 mM $CaCl_2$, and 10% glycerol and then loaded onto an anion-exchange column (100 × 10 mm i.d., Waters Accell Plus QMA, 500-Å pore size, 50-μm average particle size). The column was developed with a linear NaCl gradient at a change of 32 mM NaCl/min and a flow rate of 2 mL/min, and the protein eluted at 280 mM NaCl. Typical yields were 20 mg of protein per 3 L of cells.

At this point, the protein was greater than 95% pure as estimated by Coomassie blue stained SDS–PAGE gels and reversed-phase HPLC. Reversed-phase HPLC was done on a Vydac C18 column (250 × 10 mm i.d., 500-Å pore size, 10-μm particle size, Separations Group). The protein was loaded on the column at 10% acetonitrile and 0.1% trifluoroacetic acid in an aqueous solution. The column was developed at 2.3 mL/min with a linear gradient of 1% acetonitrile/min, and the protein eluted at 49% acetonitrile. This protein was used in electrospray ionization mass spectroscopy in order to confirm the amino acid composition. For both Sc-t96 and Kl-t96, protein concentrations were determined using an extinction coefficient of 7000 OD_{280} M^{-1} cm^{-1}, which was calculated from the numbers of tyrosines and tryptophans in the two proteins.

Circular Dichroism Spectroscopy. Circular dichroism (CD) spectra were recorded on an AVIV 60 DS spectropolarimeter at 25 °C. Measurements were taken at 1-nm intervals with a 1-s time constant and a 1.5-nm bandwidth, and were averaged over 5 scans. A path length of 2 mm was used for approximately 50 μM protein concentration. Unless otherwise noted, the buffer conditions were 25 mM sodium phosphate, pH 8.8, and 100 mM NaF. Fluoride was used as an anion to avoid the higher absorption of chloride ions in the far-ultraviolet range. In order to determine the fractional helicity (f_H) of the proteins, we used the mean residue ellipticity (θ) at 222 nm in the equation from Chen et al. (1972):

$$[\theta]_{222} = -30300 f_H - 2340$$

Cross-Linking Reactions. Stock solutions of ethylene glycol bis(succinimidyl succinate) (EGS) were prepared fresh in dimethyl sulfoxide, and added to a 20 μM solution of the protein in 25 mM sodium phosphate, pH 8.8, to give final EGS concentrations ranging from 0.8 to 400 μM. The reactions proceeded for 15 min at room temperature and were then quenched by addition of glycine to a final concentration of 40 mM. The products were analyzed by electrophoresis on denaturing 5–20% gradient polyacrylamide gels; bands were visualized with Coomassie blue.

Analytical Ultracentrifugation. Sedimentation equilibria and sedimentation velocities were determined with a Beckman Optima XL-A ultracentrifuge. The speeds were 15 000 and 60 000 rpm, respectively, with a radius (r) of 5.80–7.20 cm. The temperature was set to 10 °C. The starting protein concentration varied from 2 to 100 μM, in a buffer containing 25 mM sodium phosphate, pH 8.8, and 100 mM NaCl. Given these buffer conditions, we used standard tables (Weast, 1975) to calculate the solvent density (ρ) as 1.008 g mL^{-1} and the solvent viscosity (η) as 0.01335 g $(cm\ s)^{-1}$. The partial specific volumes (ν_2) for Sc-t96 and Kl-t96 were calculated as 0.7266 and 0.7333 cm^3 g^{-1}, respectively (McMeekin et al., 1975). We assumed a hydration shell of 0.34 g of H_2O/g of protein in order to calculate hydrated volumes (V_h) of 59.2 and 58.8 nm^3 for Sc-t96 and Kl-t96, respectively.

Data analysis was done with programs supplied by Beckman. The sedimentation equilibrium data were evaluated assuming a two-component system (solvent and solute), fitting a model curve to the observed distribution (Van Holde, 1975). While this fits extremely well for Kl-t96, it does not fit as well for Sc-t96 at higher protein concentrations due to the formation of larger complexes at the high end of the gradient (see below). For measurements of sedimentation velocity, the inflection point of the boundary was used (Goldberg, 1953).

Light Scattering. The diffusion coefficient was measured by dynamic light scattering using a Biotage dp-801 detector (Biotage, Inc., Charlottesville, VA). The protein concentration was 2 mg mL^{-1} in a buffer containing 25 mM sodium phosphate, pH 8.8, and 100 mM NaCl. The solution was filtered through a 20-nm syringe filter (Anotec Separations, Ltd.). Measurements were done at 25 °C. Given these buffer conditions, we used standard tables (Weast, 1975) to calculate the solvent viscosity (η) as 0.009065 g $(cm\ s)^{-1}$. To calculate the hydrodynamic radius (R_h), we used the following equation: $R_h = kT/6\pi\eta D$, where k is Boltzmann's constant and T is the absolute temperature. The hydrodynamic radius was used as the long semi-axis dimension (a) in order to calculate the short semi-axis dimension (b) of a prolate ellipsoid with a hydrated volume $V_h = (4/3)\pi ab^2$.

RESULTS

Overexpression, Purification, and Characterization. Fragments of the *S. cerevisiae* and *K. lactis* HSF genes containing the putative trimerization domain were overexpressed in *E. coli*. The predicted amino acid sequences of the Sc-t96 and Kl-t96 proteins, including the amino acids introduced during the cloning procedure, are shown in Figure 1. Note that Sc-t96 and Kl-t96 contain 92 and 91 naturally occurring amino acids, respectively, and Kl-t96 defines a new N-terminal boundary for the putative trimerization domain. Electrospray ionization mass spectroscopy of the Sc-t96 and Kl-t96 proteins expressed in *E. coli* shows predominant species with molecular weights of 11 137.17 ± 1.03 and 10 980.61 ± 0.85, respectively. This agrees extremely well with the predicted molecular weights of 11 137.81 and 10 980.31 for Sc-t96 and Kl-t96 calculated from their amino acid sequence, indicating that there is no posttranslational modification or processing of these proteins in *E. coli*.

```
                    *   -   *   -   *   -   *
Sc-t96: MNEANLGDKNDVTAILGELEQIKYNQIAISKDLLR
Kl-t96: mLGGTTNAEVDIHILLNELETVKYNQLAIAEDLKR

        -   *   -   *   -
        INKDNELLWQENMMARERHRTQQQALEKNRFFLTS
        ITKDNEMLWKENMMARERHQSQQQVLEKLLRFLSS

        IVPHLDPKMIMDGLGDPKVNNEgmln
        VFGPNSAKTIGNGFGPDLIHELgmln
```

FIGURE 1: Sequence of Sc-t96 and Kl-t96. The sequences for the two proteins are shown above. The residues that could potentially form part of the heptad array are indicated by a star and a dash. The residues highly conserved between all sequenced trimerization domains are in boldface print. The residues predicted to be α-helical by Chou–Fasman rules are underlined. The residues in lower-case were introduced during the cloning procedure (see Materials and Methods).

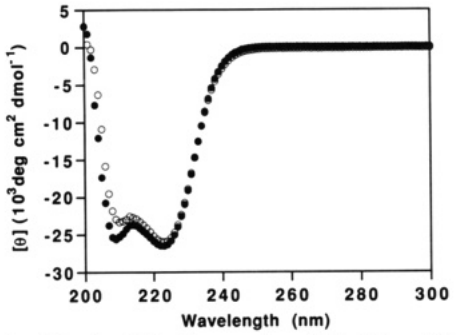

FIGURE 2: Circular dichroism spectra of Sc-t96 and Kl-t96 in 25 mM sodium phosphate, pH 8.8, and 100 mM NaF. Five scans were averaged for each CD spectrum. Open circules (○) are Sc-t96 (measured at 51.8 μM); solid circles (●) are Kl-t96 (measured at 48.2 μM).

Secondary Structure. We have analyzed the secondary structure of these two proteins using circular dichroism. As shown in Figure 2, both proteins have CD spectra characteristic of α-helical structures, with minima at 222 and 208 nm. The α-helical content of both proteins is independent of protein concentration in the range from 1 to 100 μM (data not shown). In addition, the spectra were independent of pH in the range from 6 to 11, anion (fluoride versus chloride) or NaCl concentration in the range of 20 mM–1 M NaCl (data not shown). The fact that the α-helical content is essentially invariant over a wide range of conditions indicates that this domain has a stable secondary structure. Using the mean residue ellipticity at 222 nm of –25 842 cm^2 deg dmol^{-1} for Sc-t96 and –26 439 cm^2 deg dmol^{-1} for Kl-t96, we estimate the fractional helical content as 78% and 80%, respectively. These values are consistent with predictions of 68% and 79%, respectively, on the basis of the primary sequence (Chou & Fasman, 1978).

Quaternary Structure by Cross-Linking. To determine the quaternary structure of these proteins in solution, we treated the proteins for a limited time with different amounts of EGS, a homobifunctional cross-linking agent. The major species observed in this experiment were complexes which ran on a SDS–PAGE gel with mobilities comparable to monomers, dimers, and trimers of the 96 amino acid proteins (see Figure 3). When we repeated this experiment at high protein (100 μM) and salt (100 mM) concentrations, we observed some larger complexes for Sc-t96 at the highest EGS concentrations (2 mM; data not shown). Interestingly, the formation of these larger complexes occurs under conditions which do not affect secondary structure as seen by the CD spectrum. We suspect that these larger complexes are aggregates of trimers.

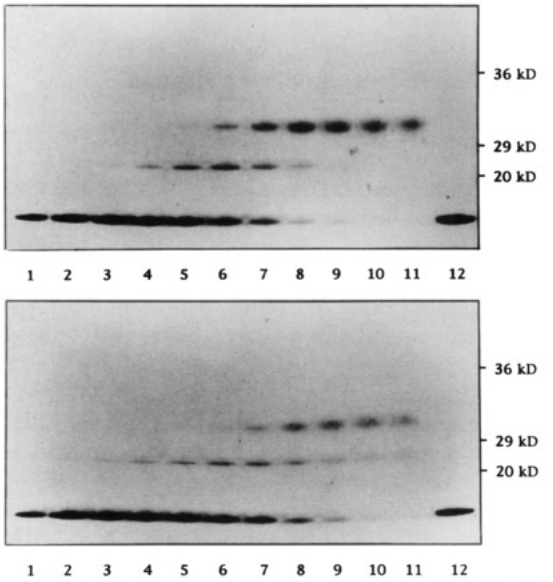

FIGURE 3: Chemically cross-linked Sc-t96 and Kl-t96 on denaturing 5–20% gradient polyacrylamide gels. Sc-t96 (top) and Kl-t96 (bottom) were cross-linked as described under Materials and Methods. Lanes 1–11 contain 0, 0.8, 1.6, 3.1, 6.2, 12.5, 25, 50, 100, 200, and 400 μM EGS, respectively. Lane 12 contains protein alone.

Quaternary Structure by Analytical Ultracentrifugation. We used analytical ultracentrifugation as an additional approach to determine the quaternary structure of these 96 amino acid proteins in solution. For Kl-t96, the molecular weight as determined by sedimentation equilibrium averaged over all initial protein loads of 2–100 μM was 30.47K ± 1.37K, which is in good agreement with the predicted molecular weight of 32.9K for a trimeric complex. For Sc-t96, the molecular weight as determined by sedimentation equilibrium at low loading protein concentrations (2 μM) was 34.8K, which agrees well with the predicted molecular weight of 33.4K for the trimeric complex. However, at higher loading protein concentrations, the slope of the gradient did not fit well to a two-component system, assuming solvent and trimeric complex as the two components. The slope of the gradient at the bottom end of the cell was steeper than expected, indicating the prevalence of larger molecular weight complexes at extremely high protein concentrations. One explanation could be the formation of oligomers, such as dimers and trimers, of the trimeric complex.

We measured sedimentation coefficients from sedimentation velocity experiments for Sc-t96 and Kl-t96 from 2 to 100 μM (Figure 4). The sedimentation coefficients were approximately constant for loading protein concentrations from 2 to 50 μM. In this range, the average values for $s_{20,w}$ for Sc-t96 and Kl-t96 were 2.6 and 2.5 S, respectively. If we use calculated hydration volumes and assume a spherical shape for the complexes, we would have predicted sedimentation coefficients of 3.3 and 3.2 S, respectively. The fact that the measured sedimentation coefficients are smaller indicates that the trimeric complexes have an elongated, nonspherical shape. This agrees with results from gel filtration experiments: Sc-t96 and Kl-t96 elute more rapidly from the column than predicted from the standard elution volumes measured with globular proteins. These results also imply that Sc-t96 and Kl-t96 are more rodlike than spherical in shape (data not shown).

Dynamic Light Scattering. We used dynamic light scattering to measure the translational diffusion coefficients of two proteins. Sc-t96 was too polydisperse to give accurate diffusion coefficients. This could be explained by the fact

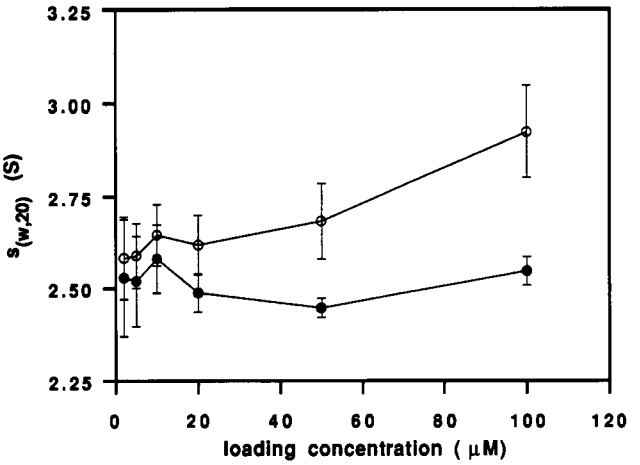

FIGURE 4: Sedimentation velocity. Sedimentation coefficients vs protein concentration for Sc-t96 (○) and Kl-t96 (●).

that Sc-t96 trimers form larger complexes (see above) under the high protein concentrations (180 μM) necessary for the dynamic light-scattering measurements. However, Kl-t96 was monodisperse, and gave a diffusion coefficient ($D_{20,w}$) of 6.48 × 10^{-7} cm^2 s^{-1}. Using the diffusion and sedimentation coefficients, we calculate the molecular weight as 34.5K. This is in good agreement with the equilibrium sedimentation results (30.5K), as well as the predicted molecular weight of a trimeric complex (32.9K).

From the measured diffusion coefficient, we calculate a hydrodynamic radius of 3.3 nm. This value differs from the calculated hydrodynamic radius of 2.4 nm, assuming a spherical shape. This is not too surprising given the results from the velocity sedimentation experiments. If we assume that the trimeric complex of Kl-t96 has a prolate ellipsoid shape, we can calculate that the other semi-axis dimension is 2.1 nm. The implications of these results on the shape for the complexes are discussed below.

DISCUSSION

In this paper, we present experimental evidence indicating that 91 amino acids from the HSF protein contain all the structural elements necessary for folding into a trimerization domain. The high α-helical content predicted from the primary structure and the heptad repeats is shown to be present by CD spectroscopy. The oligomerization state in solution was determined by cross-linking and sedimentation equilibrium studies. In addition, sedimentation velocity and light-scattering measurements are consistent with an elongated, nonspherical shape. These approaches confirm that this domain forms a trimeric complex, with a high α-helical content, suggesting an α-helical coiled-coil as a central motif.

While the fragment from *K. lactis* (Kl-t96) behaves as an ideal trimeric complex, the fragment from *S. cerevisiae* (Sc-t96) forms larger complexes at high protein concentration (≥100 μM), as demonstrated by equilibrium sedimentation. We suspect that these larger species are due to aggregation between trimers because they can be eliminated, without affecting the trimeric complexes, by reducing the salt concentration. In addition, dynamic light scattering at high protein (180 μM) and salt (100 mM NaCl) concentrations indicates that the Sc-t96 protein is slightly polydisperse under these conditions, while Kl-t96 is monodisperse (data not shown). The fact that Kl-t96 does not form complexes larger than trimers in the same concentration range as Sc-t96 supports our conclusion that the observed larger complexes of Sc-t96 are an artifact of certain in vitro conditions and that both proteins preferentially and primarily form trimeric complexes.

```
Sce  (333-424)  MNEANLGDKNDVTAILGELEQIKYNQIAISKDLLRINKDNELL
Kla  (303-394)  ILGGTTNAEVDIHILLNELETVKYNQLAIAEDLKRITKDNEML
Dro  (155-246)  GDDKGVLKPEAMSKILTDVKVMRGRQDNLDSFRSAMKQENEVL
Mo1  (126-217)  KSEDIKIRQDSVTRLLTDVQLMKGKQECMDSKLLAMKHENEAL
Mo2  (115-206)  KPEENKIRQEDLTKIISSAQKVQIKQETIESRLSELKSENESL
Hu1  (126-217)  KSEDIKIRQDSVTKLLTDVQLMKGKQECMDSKLLAMKHENEAL
Hu2  (115-206)  KPEENKIRQEDLTKIISSAQKVQIKQETIESRLSELKSENESL

Sce  WQENMMARERHRTQQQALEKMFRFLTSIVPHLDPKMIMDGLGDPKVNNE
Kla  WKENMMARERHQSQQQVLEKLLRFLSSVFGPNSAKTIGNGFQPDLIHEL
Dro  WREIASLRQKHAKQQQIVNKLIQFLITIVQPSRNMSGVKRHVQLMINNT
Mo1  WREVASLRQKHAQQQKVVNKLIQFLISLVQSNRILGVKRKIPLMLSDSN
Mo2  WKEVSELRAKHAQQQQVIRKIVQFIVTLVQNNQLVSLKRKRPLLLNTNG
Hu1  QREVASLRQKHAQQQKVVNKLIQFLISLVQSNRILGVKRKIPLMLNDSG
Hu2  WKEVSELRAKHAQQQQVIRKIVQFIVTLVQNNQLVSLKRKRPLLLNTNG
```

FIGURE 5: Sequence comparisons of trimerization domains. Sce, *S. cerevisiae* aa 333–424 (Sorger & Pelham, 1989); Kla, *K. lactis* aa 303–394 (Jakobsen & Pelham, 1991); Dro, *D. Melanogaster* aa 155–246 (Clos et al., 1990); Mo1, mouse aa 126–217 (Sarge et al., 1991); Mo2, mouse aa 115–206 (Sarge et al., 1991); Hu1, human aa 126–217 (Rabindran et al., 1991); Hu2, human aa 115–206 (Schuetz et al., 1991).

Given the α-helical and trimeric nature of these protein fragments, we postulate that the trimeric complexes form a triple-stranded α-helical coiled-coil. Using the hydrodynamic radii measured by dynamic light scattering, we can calculate overall dimensions for a prolate ellipsoid shape of 6.6 by 4.2 nm for Kl-t96. These dimensions are less elongated and more compact than would be expected for a single, long α-helical coiled-coil, as is seen for leucine zippers (O'Shea et al., 1991). Instead, we propose that HSF forms a trimeric complex analogous to that found in the hemagglutinin of influenza virus, whose structure has been solved by X-ray crystallography (Wilson et al., 1981). In the hemagglutinin trimerization domain, there are two helices: the longer helix forms a triple-stranded coiled-coil, while the shorter helix forms an outer sandwich that buttresses against the inner coiled-coil. The absolute dimensions of the hemagglutinin trimerization domain are similar to the calculated values for the HSF fragment. The hemagglutinin domain is longer than the HSF fragment (8.2 vs 6.6 nm), indicating that hemagglutinin probably has a longer central coiled-coil. In fact, the hemagglutinin coiled-coil extends for 53 amino acids, compared to the prediction of 42 amino acids (6 heptad arrays) for HSF. The X-ray structure of hemagglutinin shows a variation in the width of the short dimension from 3.0 to 4.8 nm, similar to the calculated width of 4.2 nm for HSF. This suggests that HSF also has an additional helix that buttresses the central coiled-coil core of the trimerization domain, much like that found in hemagglutinin. The existence of this postulated additional helix is supported by the fact that residues between 393 and 424 appear to be required for trimerization (Sorger & Nelson, 1989).

Is this trimerization motif common to all HSFs? As shown in Figure 5, there is significant homology between all eukaryotic HSFs sequenced to date that extends well beyond the simple heptad array. Recent results on full-length *Drosophila* HSF confirm that it is a trimeric protein (Rabindran et al., 1992). This leads us to believe that all HSFs might have a similar trimerization domain. Why is HSF trimeric? There is probably a connection between the trimerization of HSF and its ability to recognize its DNA binding site. The minimum requirement of three nGAAn boxes, as well as the alternating orientation of the five base pair units, places the contact sites of all three nGAAn boxes on the same side of the DNA double helix, in a distorted 3-fold symmetry (Xiao et al., 1991). A second plausible hypothesis is that the trimerization of HSF might be involved in regulation of the heat shock response via interactions with the heat shock protein hsp70 (Clos et al., 1990). It is intriguing

to note that hsp70 and its cognates often interact with proteins that trimerize through amphiphilic helices, such as influenza hemagglutinin (Gething et al., 1986) and clathrin (DeLuca-Flaherty et al., 1990; Näthke et al., 1992).

ACKNOWLEDGMENT

We thank the following people: Drs. Dave King, Francis Bitsch, and Cedric Shackleton, for the electrospray ionization mass spectroscopy; Dr. Susan Marqusee, for help with circular dichroism; Ying Yang and Dr. Howard Schachman, for help with the analytical ultracentrifuge; Dan Snyder (Biotage, Inc.), for help with the Biotage dp-801 detector; Drs. Alex Glazer, Susan Marqusee, and Dave King, for reading the manuscript; Eva Grotkopp, for technical assistance; and members of the Nelson laboratory, for helpful advice.

REFERENCES

Amin, J., Anathan, J., & Voellmy, R. (1988) *Mol. Cell. Biol. 8*, 3761–3769.

Chen, Y.-H., Yang, J. T., Martinez, H. M. (1972) *Biochemistry 11*, 4120–4131.

Chou, P. Y., & Fasman, G. D. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol. 47*, 45–148.

Clos, J., Westwood, J. T., Becker, P. B., Wilson, S., Lambert, K., & Wu, C. (1990) *Cell 63*, 1085–1097.

Crick, F. H. C. (1953) *Acta. Crystallogr. 6*, 689–690.

DeLuca-Flaherty, C., McKay, D. B., Parham, P., & Hill, B. L. (1990) *Cell 62*, 875–887.

Gething, M. J., McCammon, K., & Sambrook, J. (1986) *Cell 46*, 939–950.

Goldberg, R. J. (1953) *J. Phys. Chem. 57*, 194–202.

Hodges, R. S., Sodek, J., Smillie, L. B., & Jurasek, L. (1972) *Cold Spring Harbor Symp. Quant. Biol. 37*, 299–310.

Jakobsen, B., & Pelham, H. R. B. (1991) *EMBO J. 10*, 369–375.

McMeekin, T. L., Groves, M. L., & Hipp, N. J. (1949) *J. Am. Chem. Soc. 71*, 3298–3300.

Näthke, I. D., Heuser, J., Lupas, A., Stock, J., Turck, D. W., & Brodsky, F. M. (1992) *Cell 68*, 899–910.

O'Shea, E. K., Klemm, J. D., Kim, P. S., & Alber, T. (1991) *Science 254*, 539–544.

Perisic, O., Xiao, J., & Lis, J. T. (1989) *Cell 59*, 797–806.

Rabindran, S. K., Giorgi, G., Clos, J., & Wu, C. (1991) *Proc. Natl. Acad. Sci. U.S.A.* 88, 6906–6910.

Rabindran, S. K., Haroun, R. I., Clos, J., Wisniewski, J., & Wu, C. (1992) *Science* (in press).

Sarge, K. D., Zimarino, V., Holm, K., Wu, C., & Morimoto, R. I. (1991) *Genes Dev. 5*, 1902–1911.

Scharf, K.-D., Rose, S., Zott, W., Schoff, F., & Nover, L. (1990) *EMBO J. 9*, 4495–4501.

Schuetz, T. J., Gallo, G. J., Sheldon, L., Tempst, P., & Kingston, R. E. (1991) *Proc. Natl. Acad. Sci. U.S.A. 88*, 6911–6915.

Sodek, J., Hodges, R. S., Smillie, L. B., & Jurasek, L. (1972) *Proc. Natl. Acad. Sci. U.S.A. 69*, 3800–3804.

Sorger, P. K., & Pelham, H. R. B. (1988) *Cell 54*, 855–864.

Sorger, P. K., & Nelson, H. C. M. (1989) *Cell 59*, 807–813.

Studier, F. W., Rosenberg, A. H., Dunn, J. J., & Dubendorff, J. W. (1990) *Methods Enzymol. 185*, 60–89.

Van Holde, K. E. (1975) *Proteins (3rd Ed.) 1*, 225–291.

Weast, R. C. (1975) *Handbook of Chemistry and Physics*, 59th ed., CRC Press, Cleveland.

Wiederrecht, G., Seto, D., & Parker, C. S. (1988) *Cell 54*, 841–853.

Wilson, I. A., Skehel, J. J., & Wiley, D. C. (1981) *Nature 289*, 366–373.

Xiao, H., & Lis, J. T. (1988) *Science 239*, 1139–1142.

Xiao, H., Perisic, O., & Lis, J. T. (1991) *Cell 64*, 585–593.